Entropy **2014**, *xx*, 1-x; doi:10.3390/——

OPEN ACCESS
entropy
ISSN 1099-4300
www.mdpi.com/journal/entropy

Article

The entropy-based quantum metric

Roger Balian *

Institut de Physique Théorique
CEA Saclay
91191 Gif-sur-Yvette Cedex

* Author to whom correspondence should be addressed; roger@balian.fr

Received: xx / Accepted: xx / Published: xx

Abstract: The von Neumann entropy $S(\hat{D})$ generates in the space of quantum density matrices $\hat{D}$ the Riemannian metric $\mathrm{d}s^2 = -\mathrm{d}^2S(\hat{D})$, which is physically founded and which characterises the amount of quantum information lost by mixing $\hat{D}$ and $\hat{D} + \mathrm{d}\hat{D}$. A rich geometric structure is thereby implemented in quantum mechanics. It includes a canonical mapping between the spaces of states and of observables, which involves the Legendre transform of $S(\hat{D})$. The Kubo scalar product is recovered within the space of observables. Applications are given to equilibrium and non equilibrium quantum statistical mechanics. There the formalism is specialised to the relevant space of observables and to the associated reduced states issued from the maximum entropy criterion, which result from the exact states through an orthogonal projection. Von Neumann's entropy specialises into a relevant entropy. Comparison is made with other metrics. The Riemannian properties of the metric $\mathrm{d}s^2 = -\mathrm{d}^2S(\hat{D})$ are derived. The curvature arises from the non-Abelian nature of quantum mechanics; its general expression and its explicit form for q-bits are given.

Keywords: quantum entropy; metric; q-bit; information; geometry; relevant entropy

---

1. A physical metric for quantum states.

Quantum physical quantities pertaining to a given system, termed as "observables" $\hat{O}$, behave as non-commutative random variables and are elements of a C*-algebra. We will consider below systems for which these observables can be represented by n-dimensional Hermitean matrices in a finite Hilbert space $\mathcal{H}$. In quantum (statistical) mechanics, the "state" of such a system encompasses the expectation values of all its observables [1]. It is represented by a density matrix $\hat{D}$, which plays the rôle of a probability distribution, and from which one can derive the expectation value of $\hat{O}$ in the form

$$<\hat{O}>= \mathrm{Tr}\,\hat{D}\hat{O} = (\hat{D};\hat{O})\,. \tag{1}$$

Density matrices should be Hermitean ($<\hat{O}>$ is real for $\hat{O} = \hat{O}^{\dagger}$), normalised (the expectation value of the unit observable is $\mathrm{Tr}\,\hat{D} = 1$) and non-negative (variances $<\hat{O}^2> - <\hat{O}>^2$ are non-negative). They depend on $n^2 - 1$ real parameters. If we keep aside the multiplicative structure of the set of operators and focus on their linear vector space structure, Eq. (1) appears as a linear mapping of the space of observables onto real numbers. We can therefore regard the observables and the density operators $\hat{D}$ as elements of two dual vector spaces, and expectation values (1) appear as scalar products.

It is of interest to define a metric in the space of states. For instance, the distance between an exact state $\hat{D}$ and an approximation $\hat{D}_{\mathrm{app}}$ would then characterise the quality of this approximation. However, all physical quantities come out in the form (1) which lies astride the two dual spaces of observables and states. In order to build a metric having physical relevance, we need to rely on another meaningful quantity which pertains only to the space of states.

We note at this point that quantum states are probabilistic objects that gather information about the considered system. Then, the amount of missing information is measured by von Neumann's entropy

$$S(\hat{D}) \equiv -\,\mathrm{Tr}\,\hat{D}\ln\hat{D}\,. \tag{2}$$

Introduced in the context of quantum measurements, this quantity is identified with the thermodynamic entropy when $\hat{D}$ is an equilibrium state. In non-equilibrium statistical mechanics, it encompasses, in the form of "relevant entropy" (see Sec. 5 below), various entropies defined through the maximum entropy criterion. It is also introduced in quantum computation. Alternative entropies have been introduced in the literature, but they do not present all the distinctive and natural features of von Neumann's entropy, such as additivity and concavity.

As $S(\hat{D})$ is a concave function, and as it is the sole physically meaningful quantity apart from expectation values, it is natural to rely on it for our purpose. We thus define [2] the distance $\mathrm{d}s$ between two neighbouring density matrices $\hat{D}$ and $\hat{D} + \mathrm{d}\hat{D}$ as the square root of

$$\mathrm{d}s^2 = -\mathrm{d}^2 S(\hat{D}) = \mathrm{Tr}\,\mathrm{d}\hat{D}\mathrm{d}\ln\hat{D}\,. \tag{3}$$

This Riemannian metric is of the Hessian form since the metric tensor is generated by taking second derivatives of the function $S(\hat{D})$ with respect to the $n^2 - 1$ coordinates of $\hat{D}$. We may take for such coordinates the real and imaginary parts of the matrix elements, or equivalently (Sec. 6) some linear transform of these (keeping aside the norm $\mathrm{Tr}\,\hat{D} = 1$).

2. Interpretation in the context of quantum information.

The simplest example, related to quantum information theory, is that of a q-bit (two-level system or spin $\frac{1}{2}$) for which $n = 2$. Its states, represented by 2×2 Hermitean normalised density matrices, can conveniently be parameterised by the components $r_\mu = D_{12} + D_{21}$, $i(D_{12} - D_{21})$, $D_{11} - D_{22}$ ($\mu = 1, 2, 3$)

of a 3-dimensional vector $\mathbf{r}$ lying within the unit Poincaré–Bloch sphere ($r \leq 1$). From the corresponding entropy

$$S = \frac{1+r}{2} \ln \frac{2}{1+r} + \frac{1-r}{2} \ln \frac{2}{1-r} \,, \tag{4}$$

we derive the metric

$$\mathrm{d}s^2 = \frac{1}{1-r^2} \left(\frac{\mathbf{r} \cdot \mathrm{d}\mathbf{r}}{r} \right)^2 + \frac{1}{2r} \ln \frac{1+r}{1-r} \left(\frac{\mathbf{r} \times \mathrm{d}\mathbf{r}}{r} \right)^2 \,, \tag{5}$$

which is a natural Riemannian metric for q-bits, or more generally for positive 2×2 matrices.

Identifying von Neumann's entropy to a measure of missing information, we can give a simple interpretation to the distance between two states. Indeed, the concavity of entropy expresses that some information is lost when two statistical ensembles described by different density operators merge. By mixing two equal size populations described by the neighbouring distributions $\hat{D}$ and $\hat{D}' = \hat{D} + \mathrm{d}\hat{D}$ separated by a distance $\mathrm{d}s$, we lose an amount of information given by

$$\Delta S \equiv S \left(\frac{\hat{D} + \hat{D}'}{2} \right) - \frac{S(\hat{D}) + S(\hat{D}')}{2} \sim \frac{\mathrm{d}s^2}{8} \,, \tag{6}$$

and thereby directly related to the distance $\mathrm{d}s$ defined by (3).

Such a loss is very large when $\hat{D}$ and $\hat{D}'$ lie in the vicinity of a pure state (for which $S = 0$), due to the singularity of the entropy (2) and of the metric (3) at the edge of the domain of $\hat{D}$; this entails a divergence exemplified for $n = 2$ by the expression (5) when $r \to 1$. In particular, if $\hat{D}$ and $\hat{D}'$ are pure, the above loss of information ΔS behaves as $\alpha \mid \ln \alpha \mid$, where $\alpha = 1/8 \, \mathrm{Tr}(\mathrm{d}\hat{D})^2$, whereas it is of order α if they have no vanishing eigenvalue. At the other extreme, around the most disordered state $\hat{D} = \hat{I}/n$, in the region $\| \, n\hat{D} - \hat{I} \, \| \ll 1$, the metric becomes Euclidean since $\mathrm{d}s^2 \sim n \, \mathrm{Tr}(\mathrm{d}\hat{D})^2$ (for $n = 2$, $\mathrm{d}s^2 = \mathrm{d}\mathbf{r}^2$). For a given shift $\mathrm{d}\hat{D}$, the qualitative change of a state $\hat{D}$, as measured by the distance $\mathrm{d}s$, gets larger and larger as the state $\hat{D}$ becomes purer and purer, that is, when the information contents of $\hat{D}$ increases.

3. Geometry of quantum statistical mechanics.

A rich geometric structure is generated for both states and observables by von Neumann's entropy through introduction of the metric $\mathrm{d}s^2 = -\mathrm{d}^2 S$. Now, this metric (3) supplements the algebraic structure of the set of observables and the above duality between the vector spaces of states and of observables, with scalar product (1). Accordingly, we can define naturally within the space of states scalar products, geodesics, angles, curvatures.

We can also regard the coordinates of $\mathrm{d}\hat{D}$ and $\mathrm{d} \ln \hat{D}$ as covariant and contravariant components of the same infinitesimal vector (Sec. 6). To this aim, let us introduce the mapping

$$\hat{D} \equiv \frac{e^{\hat{X}}}{\mathrm{Tr} \, e^{\hat{X}}} \tag{7}$$

between $\hat{D}$ in the space of states and $\hat{X}$ in the space of observables. The operator $\hat{X}$ appears as a parameterisation of $\hat{D}$. (The normalisation of $\hat{D}$ entails that $\hat{X}$, defined within an arbitrary additive constant operator $X_0\,\hat{I}$, also depends on n^2-1 independent real parameters.) The metric (3) can then be re-expressed in terms of $\hat{X}$ in the form

$$\mathrm{d}s^2 = \mathrm{Tr}\,\mathrm{d}\hat{D}\mathrm{d}\hat{X} = \mathrm{Tr}\int_0^1 \mathrm{d}\xi\,\hat{D}e^{-\xi\hat{X}}\mathrm{d}\hat{X}e^{\xi\hat{X}}\mathrm{d}\hat{X} - (\mathrm{Tr}\,\hat{D}\mathrm{d}\hat{X})^2 = \mathrm{d}^2\ln\mathrm{Tr}\,e^{\hat{X}} = \mathrm{d}^2F\,, \tag{8}$$

where we introduced the function

$$F(\hat{X}) \equiv \ln\mathrm{Tr}\,e^{\hat{X}} \tag{9}$$

of the observable $\hat{X}$. (The addition of $X_0\hat{I}$ to $\hat{X}$ results in the addition of the irrelevant constant X_0 to F.) This mapping provides us with a natural metric in the space of observables, from which we recover the scalar product between $\mathrm{d}\hat{X}_1$ and $\mathrm{d}\hat{X}_2$ in the form of a Kubo correlation in the state $\hat{D}$. The metric (8) has been quoted in the literature under the names of Bogoliubov–Kubo–Mori.

4. Covariance and Legendre transformation.

We can recover the above geometric mapping (7) between $\hat{D}$ and $\hat{X}$, or between the covariant and contravariant coordinates of $\mathrm{d}\hat{D}$, as the outcome of a Legendre transformation, by considering the function $F(\hat{X})$. Taking its differential $\mathrm{d}F = \mathrm{Tr}\,e^{\hat{X}}\mathrm{d}\hat{X}/\,\mathrm{Tr}\,e^{\hat{X}}$, we identify the partial derivatives of $F(\hat{X})$ with the coordinates of the state $\hat{D} = e^{\hat{X}}/\,\mathrm{Tr}\,e^{\hat{X}}$, so that $\hat{D}$ appears as conjugate to $\hat{X}$ in the sense of Legendre transformations. Expressing then $\hat{X}$ as function of $\hat{D}$ and inserting into $F - \mathrm{Tr}\,\hat{D}\hat{X}$, we recognise that the Legendre transform of $F(\hat{X})$ is von Neumann's entropy $F - \mathrm{Tr}\,\hat{D}\hat{X} = S(\hat{D}) = -\,\mathrm{Tr}\,\hat{D}\ln\hat{D}$. The conjugation between $\hat{D}$ and $\hat{X}$ is embedded in the equations

$$\mathrm{d}F = \mathrm{Tr}\,\hat{D}\mathrm{d}\hat{X}\ ;\quad \mathrm{d}S = -\,\mathrm{Tr}\,\hat{X}\mathrm{d}\hat{D}\,. \tag{10}$$

Legendre transformations are currently used in equilibrium thermodynamics. Let us show that they come out in this context directly as a special case of the present general formalism. The entropy of thermodynamics is a function of the extensive variables, energy, volume, particle numbers, etc. Let us focus for illustration on the energy U, keeping the other extensive variables fixed. The thermodynamic entropy $S(U)$, a function of the single variable U, generates the inverse temperature as $\beta = \partial S/\partial U$. Its Legendre transform is the Massieu potential $F(\beta) = S - \beta U$. In order to compare these properties with the present formalism, we recall how thermodynamics comes out in the framework of statistical mechanics. The thermodynamic entropy $S(U)$ is identified with the von Neumann entropy (2) of the Boltzmann–Gibbs canonical equilibrium state $\hat{D}$, and the internal energy with $U = \mathrm{Tr}\,\hat{D}\hat{H}$. In the relation (7), the operator $\hat{X}$ reads $\hat{X} = -\beta\hat{H}$ (within an irrelevant additive constant). By letting U or β vary, we select within the spaces of states and of observables a one-dimensional subset. In these restricted subsets, $\hat{D}$ is parameterised by the single coordinate U, and the corresponding $\hat{X}$ by the coordinate $-\beta$. By specialising the general relations (10) to these subsets, we recover the thermodynamic relations $\mathrm{d}F = -U\mathrm{d}\beta$ and $\mathrm{d}S = \beta\mathrm{d}U$. We also recover, by restricting the metric (3) or (8) to these subsets, the current thermodynamic metric $\mathrm{d}s^2 = -(\partial^2S/\partial U^2)\mathrm{d}U^2 = -\mathrm{d}U\mathrm{d}\beta$.

More generally, we can consider the Boltzmann–Gibbs states of equilibrium statistical mechanics as the points of a manifold embedded in the full space of states. The thermodynamic extensive variables, which parameterise these states, are the expectation values of the conserved macroscopic observables, that is, they are a subset of the expectation values (1) which parameterise arbitrary density operators. Then the standard geometric structure of thermodynamics simply results from the restriction of the general metric (3) to this manifold of Boltzmann–Gibbs states. The commutation of the conserved observables simplifies the reduced thermodynamic metric, which presents the same features as a Fisher metric (see Sec. 6).

5. Relevant entropy and geometry of the projection method.

The above ideas also extend to non-equilibrium quantum statistical mechanics [2,3]. When introducing the metric (3), we indicated that it may be used to estimate the quality of an approximation. Let us illustrate this point with the Nakajima–Zwanzig–Mori–Robertson projection method, best introduced through maximum entropy. Consider some set $\{\hat{A}_k\}$ of "relevant observables", whose time-dependent expectation values $a_k \equiv < \hat{A}_k > = \operatorname{Tr} \hat{D}\hat{A}_k$ we wish to follow, discarding all other variables. The exact state $\hat{D}$ encodes the variables $\{a_k\}$ that we are interested in, but also the expectation values (1) of the other observables that we wish to eliminate. This elimination is performed by associating at each time with $\hat{D}$ a "reduced state" $\hat{D}_{\mathrm{R}}$ which is equivalent to $\hat{D}$ as regards the set $a_k = \operatorname{Tr} \hat{D}_{\mathrm{R}}\hat{A}_k$, but which provides no more information than the values$\{a_k\}$. The former condition provides the constraints $< \hat{A}_k > = a_k$, and the latter condition is implemented by means of the maximum entropy criterion: One expresses that, within the set of density matrices compatible with these constraints, $\hat{D}_{\mathrm{R}}$ is the one which maximises von Neumann's entropy (2), that is, which contains solely the information about the relevant variables a_k. The least biased state $\hat{D}_{\mathrm{R}}$ thus defined has the form $\hat{D}_{\mathrm{R}} = e^{\hat{X}_{\mathrm{R}}} / \operatorname{Tr} e^{\hat{X}_{\mathrm{R}}}$, where $\hat{X}_{\mathrm{R}} \equiv \sum_k \lambda_k \hat{A}_k$ involves the time-dependent Lagrange multipliers λ_k, which are related to the set a_k through $\operatorname{Tr} \hat{D}_{\mathrm{R}}\hat{A}_k = a_k$.

The von Neumann entropy $S(\hat{D}_{\mathrm{R}}) \equiv S_{\mathrm{R}}\{a_k\}$ of this reduced state $\hat{D}_{\mathrm{R}}$ is called the "relevant entropy" associated with the considered relevant observables $\hat{A}_k$. It measures the amount of missing information, when only the values $\{a_k\}$ of the relevant variables are given. During its evolution, $\hat{D}$ keeps track of the initial information about all the variables $< \hat{O} >$ and its entropy $S(\hat{D})$ remains constant in time. It is therefore smaller than the relevant entropy $S(\hat{D}_{\mathrm{R}})$ which accounts for the loss of information about the irrelevant variables. Depending on the choice of relevant observables $\{\hat{A}_k\}$, the corresponding relevant entropies $S_{\mathrm{R}}\{a_k\}$ encompass various current entropies, such as the non-equilibrium thermodynamic entropy or Boltzmann's H-entropy.

The same structure as the one introduced above for the full spaces of observables and states is recovered in this context. Here, for arbitrary values of the parameters λ_k, the exponents $\hat{X}_{\mathrm{R}} = \sum_k \lambda_k \hat{A}_k$ constitute a subspace of the full vector space of observables, and the parameters $\{\lambda_k\}$ appear as the coordinates of $\hat{X}_{\mathrm{R}}$ on the basis $\{\hat{A}_k\}$. The corresponding states $\hat{D}_{\mathrm{R}}$, parameterised by the set $\{a_k\}$, constitute a subset of the space of states, the manifold $\mathcal{R}$ of "reduced states". (Note that this manifold is not a hyperplane, contrary to the space of relevant observables; it is embedded in the full vector space of states, but does not constitute a subspace.) By regarding $S_{\mathrm{R}}\{a_k\}$ as a function of the coordinates $\{a_k\}$, we can define a metric $\mathrm{d}s^2 = -\mathrm{d}^2 S_{\mathrm{R}}\{a_k\}$ on the manifold $\mathcal{R}$, which is the restriction of the metric (3). Its alternative expression $\mathrm{d}s^2 = \sum_k \mathrm{d}a_k \mathrm{d}\lambda_k = \mathrm{d}^2 F_{\mathrm{R}}\{\lambda_k\}$, where $F_{\mathrm{R}}\{\lambda_k\} \equiv \ln \operatorname{Tr} \exp \sum_k \lambda_k \hat{A}_k$, is a restriction of (8). The correspondence between the two parameterisations $\{a_k\}$ and $\{\lambda_k\}$ is again implemented by the Legendre transformation which relates $S_{\rm R}\{a_k\}$ and $F_{\rm R}\{\lambda_k\}$.

The projection method relies on the mapping $\hat{D} \mapsto \hat{D}_{\rm R}$ from $\hat{D}$ to $\hat{D}_{\rm R}$. It consists in replacing the Liouville–von Neumann equation of motion for $\hat{D}$ by the corresponding dynamical equation for $\hat{D}_{\rm R}$ on the manifold $\mathcal{R}$, or equivalently for the coordinates $\{a_k\}$ or for the coordinates$\{\lambda_k\}$, a programme that is in practice achieved through some approximations. This mapping is obviously a projection in the sense that $\hat{D} \mapsto \hat{D}_{\rm R} \mapsto \hat{D}_{\rm R}$, but moreover the introduction of the metric (3) shows that the vector $\hat{D} - \hat{D}_{\rm R}$ in the space of states is perpendicular to the manifold $\mathcal{R}$ at the point $\hat{D}_{\rm R}$. This property is readily shown by writing, in this metric, the scalar product $\mathrm{Tr}\, \mathrm{d}\hat{D}\mathrm{d}\hat{X}'$ of the vector $\mathrm{d}\hat{D} = \hat{D} - \hat{D}_{\rm R}$ by an arbitrary vector $\mathrm{d}\hat{D}'$ in the tangent plane of $\mathcal{R}$. The latter is conjugate to any combination $\mathrm{d}\hat{X}'$ of observables $\hat{A}_k$, and this scalar product vanishes because $\mathrm{Tr}\, \hat{D}\hat{A}_k = \mathrm{Tr}\, \hat{D}_{\rm R}\hat{A}_k$. Thus the mapping $\hat{D} \mapsto \hat{D}_{\rm R}$ appears as an orthogonal projection, so that the relevant state $\hat{D}_{\rm R}$ associated with $\hat{D}$ may be regarded as its best possible approximation on the manifold $\mathcal{R}$.

6. Properties of the metric.

The metric tensor can be evaluated explicitly in a basis where the matrix $\hat{D}$ is diagonal. Denoting by D_i its eigenvalues and by $\mathrm{d}D_{ij}$ the matrix elements of its variations, we obtain from (3)

$$\mathrm{d}s^2 = \mathrm{Tr} \int_0^\infty \mathrm{d}\xi \left(\frac{\mathrm{d}\hat{D}}{\hat{D} + \xi} \right)^2 = \sum_{ij} \frac{\ln D_i - \ln D_j}{D_i - D_j} \mathrm{d}D_{ij} \mathrm{d}D_{ji} \,. \tag{11}$$

In the same basis, the form (8) of the metric reads

$$\mathrm{d}s^2 = \frac{1}{Z} \sum_{ij} \frac{e^{X_i} - e^{X_j}}{X_i - X_j} \mathrm{d}X_{ij} \mathrm{d}X_{ji} - \left(\frac{\sum_i e^{X_i} \mathrm{d}X_{ii}}{Z} \right)^2 \,, \tag{12}$$

with $Z = \sum_i e^{X_i}$. The singularity of the metric (11) in the vicinity of vanishing eigenvalues of $\hat{D}$, in particular near pure states (end of Sec.2), is not apparent in the representation (12) of this metric, because the mapping from $\hat{D}$ to $\hat{X}$ sends the eignevalue X_i to $-\infty$ when D_i tends to zero.

Let us compare the expression (11) with the corresponding classical metric, which is obtained by starting from Shannon's entropy instead of von Neumann's entropy. For discrete probabilities p_i, we have then $S\{p_i\} = -\sum_i p_i \ln p_i$ and hence $\mathrm{d}s^2 = \sum_i \mathrm{d}p_i^2/p_i$, which is the Fisher information metric. The present metric thus appears as the extension to quantum statistical mechanics of the Fisher metric. In fact, in the terms of (11) which involve the diagonal elements $i = j$ of $\mathrm{d}\hat{D}$, the ratio reduces to $1/D_i$. This result was expected since density matrices behave as probability distributions if both $\hat{D}$ and $\mathrm{d}\hat{D}$ are diagonal. Another matrix extension of the Fisher information metric, the Bures metric

$$\mathrm{d}s^2 = \sum_{ij} \frac{2}{D_i + D_j} \mathrm{d}D_{ij} \mathrm{d}D_{ji} \,, \tag{13}$$

was introduced long ago [4], without physical justification. The elements of its metric tensor approximate the corresponding ones in (11) when $\mid D_i - D_j \mid \ll D_i + D_j$.

In order to express the properties of the Riemannian metric (3) in a general form, which will exhibit the tensor structure, we use a Liouville representation. There, the observables $\hat{O} = O_\mu \hat{\Omega}^\mu$, regarded as elements of a vector space, are represented by their coordinates O_μ on a complete basis $\hat{\Omega}^\mu$ of n^2 observables. The space of states is spanned by the dual basis $\hat{\Sigma}_\mu$, such that $\mathrm{Tr}\,\hat{\Omega}^\nu\hat{\Sigma}_\mu = \delta^\nu_\mu$, and the states $\hat{D} = D^\mu\hat{\Sigma}_\mu$ are represented by their coordinates D_μ. Thus, the expectation value (1) is the scalar product $D^\mu O_\mu$. In the matrix representation which appears as a special case, μ denotes a pair of indices $i,\ j$, $\hat{\Omega}^\mu$ stands for $\mid j >< i \mid$, $\hat{\Sigma}_\mu$ for $\mid i >< j \mid$, O_μ denotes the matrix element O_{ji} and D^μ the element D_{ij}. For the q-bit ($n = 2$) considered in Sec. 2, we have chosen the Pauli operators $\hat{\sigma}^\mu$ as basis $\hat{\Omega}^\mu$ for observables, and $1/2\hat{\sigma}_\mu$ as dual basis $\hat{\Sigma}_\mu$ for states, so that the coordinates $D^\mu = \mathrm{Tr}\,\hat{D}\hat{\Omega}^\mu$ of $\hat{D} = 1/2(\hat{I} + r^\mu\hat{\sigma}_\mu)$ are the components r^μ of the vector $\mathbf{r}$. (The unit operator $\hat{I}$ is kept aside since $\hat{D}$ is normalised and since constants added to $\hat{X}$ are irrelevant.) The function $F\{X\} = \ln \mathrm{Tr}\, e^{\hat{X}}$ of the coordinates X_μ of the observable $\hat{X}$, and the von Neumann entropy $S\{D\}$ as function of the coordinates D^μ of the state $\hat{D}$, are related by the Legendre transformation $F = S + D^\mu X_\mu$, and the relations (10) are expressed by $D^\mu = \partial F/\partial X_\mu$, $X_\mu = -\partial S/\partial D^\mu$. The metric tensor is given by

$$g^{\mu\nu} = \frac{\partial^2 F}{\partial X_\mu \partial X_\nu}, \quad g_{\mu\nu} = -\frac{\partial^2 S}{\partial D^\mu \partial D^\nu}\,, \tag{14}$$

and the correspondence issued from (7) between covariant and contravariant infinitesimal variations of $\hat{X}$ and $\hat{D}$ is implemented as $\mathrm{d}D^\mu = g^{\mu\nu}\mathrm{d}X_\nu$, $\mathrm{d}X_\mu = g_{\mu\nu}\mathrm{d}D^\nu$.

The Hessian property of the metric tensor simplifies the expression of the Christoffel symbol, which reduces to

$$\Gamma_{\mu\nu\rho} = -\frac{1}{2}\frac{\partial^3 S}{\partial D^\mu \partial D^\nu \partial D^\rho}\,. \tag{15}$$

Then, the Riemann curvature tensor comes out as

$$R_{\mu\rho\ \nu\sigma} = g^{\xi\zeta}(\Gamma_{\mu\sigma\xi}\Gamma_{\nu\rho\zeta} - \Gamma_{\mu\nu\xi}\Gamma_{\rho\sigma\zeta})\,, \tag{16}$$

the Ricci tensor and the scalar curvature as

$$R_{\mu\nu} = g^{\rho\sigma}R_{\mu\rho\ \nu\sigma}, \quad R = g^{\mu\nu}R_{\mu\nu}\,. \tag{17}$$

We have noted that the classical equivalent of the metric $\mathrm{d}s^2 = -\mathrm{d}^2 S$ is the Fisher metric $\sum_i \mathrm{d}p_i^2/p_i$, which as regards the curvature is equivalent to a Euclidean metric. While the space of classical probabilities is thus flat, the above equations show that the space of quantum states is curved. This curvature arises from the non-commutation of the observables, it vanishes for the completely disordered state $\hat{D} = \hat{I}/n$. Curvature can thus be used as a measure of the degree of classicality of a state.

In the illustrative example of a q-bit, the operator $\hat{X} = \chi_\mu\hat{\sigma}^\mu$ associated with $\hat{D}$ is parameterised by the 3 components of the vector χ_μ ($\mu = 1,\,2,\,3$), related to $\mathbf{r}$ by $\chi = \tanh^{-1} r$ and $\chi_\mu/\chi = r_\mu/r$. The metric tensor given by (5) is expressed as

$$\begin{aligned} g_{\mu\nu} &= K r_\mu r_\nu + \frac{\chi}{r}\delta_{\mu\nu}\,, \quad K \equiv \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\frac{\chi}{r} = \frac{1}{r^2}\left(\frac{1}{1-r^2} - \frac{\chi}{r}\right)\,, \\ g^{\mu\nu} &= (1-r^2)p^{\mu\nu} + \frac{r}{\chi}q^{\mu\nu}\,. \end{aligned} \tag{18}$$

(We have introduced the projectors $r^\mu r^\nu/r^2 \equiv p^{\mu\nu} \equiv \delta^{\mu\nu} - q^{\mu\nu}$ in the Euclidean 3-dimensional space so as to simplify the subsequent calculations.) We determine from (15) and (18) the explicit form

$$\Gamma_{\mu\nu\rho} = \frac{K}{2}\left(r_\mu\delta_{\nu\rho} + r_\nu\delta_{\mu\rho} + r_\rho\delta_{\mu\nu}\right) + \frac{1}{2}r\frac{\mathrm{d}K}{\mathrm{d}r}r_\mu r_\nu r_\rho \tag{19}$$

of the Christoffel symbol, then from (16) the Riemann curvature

$$R^{\mu}{}_{\rho\ \nu\sigma} = \frac{K}{4}\left[(r^2 + \frac{r}{\chi} - 1)(q^{\mu}{}_{\sigma} q_{\nu\rho} - q^{\mu}{}_{\nu} q_{\rho\sigma}) + (r^2 - \frac{r}{\chi} + 1)(p^{\mu}{}_{\sigma} q_{\nu\rho} - p^{\mu}{}_{\nu} q_{\rho\sigma}) \right. \\ \left. + \frac{r}{\chi}\frac{1}{1-r^2}(r^2 - \frac{r}{\chi} + 1)(q^{\mu}{}_{\sigma} p_{\nu\rho} - q^{\mu}{}_{\nu} p_{\rho\sigma})\right] . \tag{20}$$

Contracting with $g^{\rho\sigma}$ the indices as in (17), we finally derive the Ricci curvature

$$R^{\mu}{}_{\nu} = -\frac{Kr}{2\chi}\left(r^2 \delta^{\mu}{}_{\nu} + \frac{\chi - r}{\chi} p^{\mu}{}_{\nu}\right) , \tag{21}$$

and the scalar curvature

$$R = -\frac{Kr}{2\chi}\left(3r^2 + \frac{\chi - r}{\chi}\right) . \tag{22}$$

Both are negative in the whole Poincaré sphere. In the limit $r \to 0$, the curvature R vanishes as $R \sim -\frac{10}{9} r^2$, as expected from the general argument of Sec. 2: a weakly polarised spin behaves classically. At the other extreme $r \to 1$, R behaves as $R \sim -2\left[(1-r) \mid \ln(1-r) \mid\right]^{-1}$; it diverges, again as expected: pure states have the largest quantum nature.

The metric $\mathrm{d}s^2 = -\mathrm{d}^2 S$, introduced above in the context of quantum statistical mechanics, might more generally be useful to characterise distances in spaces of positive matrices.

References

Given the tutorial scope of this text, we refer to the articles below for extensive bibliographies.

[1] W. Thirring, *A course of mathematical physics*, vol. 4, *Quantum mechanics of large systems* (Springler Verlag, 1983)

[2] The metric $\mathrm{d}s^2 = -\mathrm{d}^2 S$ and most results reviewed here were introduced in R. Balian, Y. Alhassid and H. Reinhardt, *Dissipation in many-body systems: a geometric approach based on information theory*, Phys. Reports **131** (1986) 1-146.

[3] Simple introductions can be found in R. Balian, *Incomplete descriptions and relevant entropies*, Am. J. Phys. **67** (1989) 1078-1090 and *Information in statistical physics*, Stud. Hist. Phil. Mod. Phys. **36** (2005) 323-353

[4] D. Bures, *An extension of Kakutani's theorem*, Trans. Am. Math. Soc. **135** (1969) 199-212